Supplemental material

I. NONEQUILIBRIUM DMFT SIMULATIONS

A. Model

We consider the three-dimensional brick wall lattice illustrated in Fig. S1, which is characterized by the primitive vectors ($\mathbf{a}_1, \mathbf{a}_2$ and $\mathbf{a}_3$). The unit cell consists of two sites (A, B sublattices) and in the following we denote the position of a site in cell i and sublattice α by $\mathbf{r}_{i,\alpha}$. The position of a cell is specified by $\mathbf{r}_i = \mathbf{r}_{i,\alpha} - \mathbf{r}_{0,\alpha}$, which can be expressed as $\mathbf{r}_i = i_1\mathbf{a}_1 + i_2\mathbf{a}_2 + i_3\mathbf{a}_3$, where $i_w \in \mathbf{Z}$. The corresponding primitive vectors of the reciprocal lattice are introduced as usual, for example, $\mathbf{b}_1 = 2\pi\frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)}$. Considering a lattice system consisting of L_w unit cells

for each $\mathbf{a}_w$-direction with periodic boundary conditions in each direction, we can introduce discretized momentum vectors in the reciprocal space as

$$\mathbf{k}_l = \frac{l_1}{L_1}\mathbf{b}_1 + \frac{l_2}{L_2}\mathbf{b}_2 + \frac{l_3}{L_3}\mathbf{b}_3, \qquad (S1)$$

with $l = (l_1, l_2, l_3)$ and $l_w = 0, 1, \cdots L_w - 1$, such that $\mathbf{k}_l \cdot \mathbf{r}_i = 2\pi \sum_w \frac{i_w l_w}{L_w}$. The electron creation operators in momentum space are defined as

$$\hat{c}^\dagger_{\mathbf{k}_l,\alpha,\sigma} = \frac{1}{\sqrt{N}} \sum_i e^{i\mathbf{k}_l \cdot \mathbf{r}_i} \hat{c}^\dagger_{\mathbf{r}_i,\alpha,\sigma}, \qquad (S2)$$

where $N = L_1 L_2 L_3$ and σ the spin. With these, the kinetic part of the Hamiltonian for this lattice can be written as

$$\hat{H}_0(t) = \sum_{\mathbf{k},\sigma} \begin{bmatrix} \hat{c}^\dagger_{\mathbf{k},A,\sigma} & \hat{c}^\dagger_{\mathbf{k},B,\sigma} \end{bmatrix} \hat{h}_{\mathbf{k}}(t) \begin{bmatrix} \hat{c}_{\mathbf{k},A,\sigma} \\ \hat{c}_{\mathbf{k},B,\sigma} \end{bmatrix}, \qquad (S3a)$$

$$\hat{h}_{\mathbf{k}}(t) = \begin{bmatrix} -2J_\perp \cos(\mathbf{k} \cdot \mathbf{a}_3) & -J_{||}^*(t) - J_\perp(e^{i\mathbf{k}\cdot\mathbf{a}_2} + 1)e^{-i\mathbf{k}\cdot\mathbf{a}_1} \\ -J_{||}(t) - J_\perp(e^{-i\mathbf{k}\cdot\mathbf{a}_2} + 1)e^{i\mathbf{k}\cdot\mathbf{a}_1} & -2J_\perp \cos(\mathbf{k} \cdot \mathbf{a}_3) \end{bmatrix}. \qquad (S3b)$$

Here $J_{||} = J_x$ and $J_\perp = J_y = J_z$ in the main text. The effect of the external field along the x-axis is taken into account through the Peierls substitution

$$-J_{||}(t) = -J_{||} \exp\left[-iq\mathbf{A}(t) \cdot (\mathbf{a_1} - \tfrac{1}{2}\mathbf{a_2})\right]. \qquad (S4)$$

The vector potential $\mathbf{A}(t)$ is related to the electric field by $\mathbf{E}(t) = -\hbar\partial_t\mathbf{A}(t)$. In the present study we choose $q = 1$ and $|\mathbf{a_1} - \frac{1}{2}\mathbf{a_2}| = d_x$ and apply the field along $\mathbf{a_1} - \frac{1}{2}\mathbf{a_2}$ ($\equiv \mathbf{e}_x$). The strength and time-dependence of the field are the same as in the experiment,

$$E(t) = \frac{K_0}{d_x} \cdot \hbar\Omega \, F_{\text{ramp}}(t; t_{\text{ramp}}) \sin(\Omega t), \qquad (S5a)$$

$$F_{\text{ramp}}(t; t_{\text{ramp}}) = \begin{cases} t/t_{\text{ramp}} & t \in [0, t_{\text{ramp}}] \\ 1 & t > t_{\text{ramp}} \end{cases}. \qquad (S5b)$$

Before we proceed, we introduce a useful expression of the kinetic term. Assuming that the ramp is over or slow enough that the change of the field amplitude is negligible on the timescale of the oscillations, we can express the vector potential as $\mathbf{A}(t) = K_0 \cos(\Omega t)\mathbf{e}_x$. Expanding

Eq. (S4) into harmonics one obtains

$$\hat{H}_0(t) = -J_\perp \sum_{\langle\mathbf{i},\mathbf{j}\rangle_{y,z},\sigma} \hat{c}^\dagger_{\mathbf{i},\sigma}\hat{c}_{\mathbf{j},\sigma} - J_{||}\mathcal{J}_0(K_0) \sum_{\langle\mathbf{i},\mathbf{j}\rangle_x,\sigma} \hat{c}^\dagger_{\mathbf{i},\sigma}\hat{c}_{\mathbf{j},\sigma}$$
$$- J_{||} \sum_{\langle\mathbf{i},\mathbf{j}\rangle_x,\sigma} \sum_{n\in\mathbf{Z},n\neq 0} e^{in\frac{\pi}{2}} \mathcal{J}_n(K_0\mathbf{e}_x \cdot \mathbf{r_{ij}}) e^{in\Omega t}\hat{c}^\dagger_{\mathbf{i},\sigma}\hat{c}_{\mathbf{j},\sigma}. \qquad (S6)$$

Here $\mathbf{i}$ and $\mathbf{j}$ indicate sites regardless of the sublattice, i.e. they denote cell and sublattice indices.

The total Hamiltonian consists of $\hat{H}_0(t)$, the chemical potential term and the interaction term: $\hat{H}_{\text{tot}}(t) = \hat{H}_0(t) - \mu\hat{N} + H_{\text{int}}$. In our case, the interaction term is the local Hubbard term, Eq. (1) in the main text. In the theoretical study, we use $J_x/h = 200$ Hz as the unit of energy and choose J_y, J_z, U and Ω according to the experiment. The systems considered in this study are in the strongly-correlated regime. We solve the model with single-site DMFT, using a perturbative strong-coupling impurity solver (see the next section). It turns out that $(L_1, L_2, L_3) = (6, 6, 6)$ is large enough to reach convergence in the systems size.

B. Single-site DMFT

In order to simulate the time-evolution of the Fermi-Hubbard model on the three-dimensional brick wall lattice we employ single-site DMFT. For this, we define the
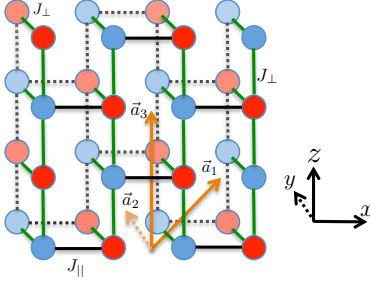
FIG. S1. Schematic picture of the three-dimensional hexagonal lattice. Red and blue circles correspond to A, B sublattices, respectively.

Green's functions in real space and momentum space as

$$G_{ij,\alpha\beta}(t,t') = -i\langle T_{\mathcal{C}}\hat{c}_{\mathbf{r}_i,\alpha,\sigma}(t)\hat{c}^\dagger_{\mathbf{r}_j,\beta,\sigma}(t')\rangle, \quad \text{(S7a)}$$

$$\hat{G}_{\mathbf{k}}(t,t') = -i\left\langle T_{\mathcal{C}}\begin{bmatrix}\hat{c}_{\mathbf{k},A,\sigma}(t)\\\hat{c}_{\mathbf{k},B,\sigma}(t)\end{bmatrix}\begin{bmatrix}\hat{c}^\dagger_{\mathbf{k},A,\sigma}(t') & \hat{c}^\dagger_{\mathbf{k},B,\sigma}(t')\end{bmatrix}\right\rangle, \quad \text{(S7b)}$$

where $\mathcal{C}$ indicates the Kadanoff-Baym (KB) contour and $T_{\mathcal{C}}$ is the contour ordering operator [S1]. Here we omitted the spin index on the left hand side assuming spin symmetry of the system. Within single-site DMFT, we assume a local self-energy on each sublattice,

$$\Sigma_{ij,\alpha\beta}(t,t') = \delta_{i,j}\delta_{\alpha,\beta}\Sigma_\alpha(t,t'). \quad \text{(S8)}$$

Under this assumption, the Green's functions in momentum space can be calculated as

$$\hat{G}_{\mathbf{k}}(t,t') = [(i\hbar\partial_t\hat{I} + \mu\hat{I} - \hat{h}_{\mathbf{k}}(t))\delta_{\mathcal{C}}(t,t') - \hat{\Sigma}(t,t')]^{-1}, \quad \text{(S9)}$$

with δ_C the delta-function on the KB contour,

$$\hat{\Sigma}(t,t') = \begin{bmatrix}\Sigma_A(t,t') & 0\\0 & \Sigma_B(t,t')\end{bmatrix},$$

and the local Green's functions is

$$G_{ii,\alpha\alpha}(t,t') = \frac{1}{N}\sum_{\mathbf{k}}G_{\mathbf{k},\alpha\alpha}(t,t'). \quad \text{(S10)}$$

In DMFT, the self-energies Σ_α are evaluated by introducing an effective impurity problem for each sublattice,

$$S_{\text{imp},\alpha} = i\int_{\mathcal{C}}dtdt'\sum_\sigma d^\dagger_\sigma(t)\mathcal{G}^{-1}_\alpha(t,t')d_\sigma(t')$$
$$- i\int_{\mathcal{C}}dtH_{\text{int}}(t), \quad \text{(S11a)}$$
$$\mathcal{G}^{-1}_\alpha(t,t') = (i\hbar\partial_t + \mu)\delta_{\mathcal{C}}(t,t') - \Delta_\alpha(t,t'). \quad \text{(S11b)}$$

Here S_{imp} is the action of the impurity problem in the path integral formalism and $d^\dagger$ indicates the Grassmann number associated with the creation operator of the electron on the impurity site. The noninteracting impurity
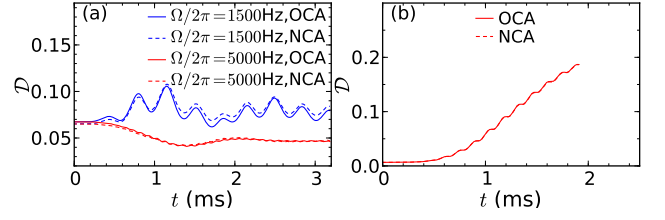

FIG. S2. Comparison of $\mathcal{D}$ obtained using NCA and OCA in the off-resonant regime (a) and the near-resonant regime (b). For the off-resonant regime, we use $J_x/h = 200$ Hz, $J_y/h = J_z/h = 40$ Hz, $U/h = 750$ Hz, $K_0 = 1.43$ and $t_{\text{ramp}} = 1.0$ ms at $T/J_x = 1.21$. For the near-resonant regime, we use $J_x/h = 200$ Hz, $J_y/h = J_z/h = 100$ Hz, $U/h = \Omega/(2\pi) = 3500$ Hz, $K_0 = 1.43$, and $t_{\text{ramp}} = 1.0$ ms at $T/J_x = 3.3$.

Green's functions $\mathcal{G}_\alpha$, or equivalently the hybridization functions Δ_α, play the role of dynamical mean fields. The interacting (spin-independent) impurity Green's function is related to the impurity self-energy $\Sigma_{\text{imp},\alpha}$ by

$$G_{\text{imp},\alpha}(t,t') = [\mathcal{G}^{-1}_\alpha(t,t') - \Sigma_{\text{imp},\alpha}(t,t')]^{-1}. \quad \text{(S12)}$$

In the DMFT self-consistency loop, the dynamical mean fields are determined such that $\Sigma_{\text{imp},\alpha} = \Sigma_\alpha$ and

$$G_{\text{imp},\alpha}(t,t') = G_{ii,\alpha\alpha}(t,t'). \quad \text{(S13)}$$

Once convergence is reached we thus obtain the lattice self-energy by calculating the impurity self-energy.

To solve the impurity problem, we mainly employ the non-crossing approximation (NCA), which is the lowest order expansion in the hybridization function and should be good in the strong coupling regime [S2]. It is known that this approximation starts to deviate from numerically exact results in the intermediate-coupling regime at low temperatures, where the one-order higher expansion (one-crossing approximation, OCA) provides important corrections [S2] at the cost of a much higher computational expense. In Fig. S2, we compare the results for $\mathcal{D}$ from NCA and OCA implementations for some representative off-resonant and near-resonant conditions. The results show that for the interactions and temperatures considered in this study, NCA and OCA yield quantitatively almost the same results (the difference is at most 5%), which justifies the use of NCA in the analysis.

C. Observables

The density of particles is calculated as

$$n(\mathbf{r}_i,t) = \frac{1}{2}\sum_{\alpha=A,B}\sum_{\sigma=\uparrow,\downarrow}\langle\hat{n}_{\mathbf{r}_{i,\alpha},\sigma}(t)\rangle. \quad \text{(S14)}$$

In the present study, we focus on the double occupation,

$$n_d(\mathbf{r}_i,t) = \frac{1}{2}\sum_{\alpha=A,B}\langle\hat{n}_{\mathbf{r}_{i,\alpha},\uparrow}(t)\hat{n}_{\mathbf{r}_{i,\alpha},\downarrow}(t)\rangle, \quad \text{(S15)}$$

which is averaged over the sublattices. In each DMFT simulation, a homogenous system is considered, hence the DMFT results show no dependence on $\mathbf{r}_i$. By running simulations for different chemical potentials at the experimentally estimated temperature (IV I), we obtain the double occupation as a function of the density of particles, which we express as $n_d[n;t]$.

In the experiment, the system is inhomogeneous because of the trapping potential, and the double-occupation ($\mathcal{D}$) is measured as the number of doubly occupied sites normalized by the total number of particles. In order to capture the effect of this inhomogeneous set-up, we employ the local density approximation. Namely, we average the results of DMFT simulations for homogeneous systems using the experimental density profile:

$$\mathcal{D}(t) = \frac{\int 2n_d[n(\mathbf{r});t]dxdydz}{\int n(\mathbf{r})dxdydz}, \qquad \text{(S16)}$$

where $\mathbf{r} = (x,y,z)$. This procedure is justified on timescales which are short enough that the redistribution of atoms in the trap can be neglected, which is the case here.

In DMFT, one can evaluate the equilibrium local spectral function

$$A(\omega) \equiv -\frac{1}{2\pi} \sum_{\alpha=A,B} \text{Im} \int dt_{\text{rel}} G^R_{ii,\alpha\alpha}(t_{\text{rel}})e^{i\omega t_{\text{rel}}} \quad \text{(S17)}$$

from the Fourier transform of the equilibrium retarded Green's functions. Here $t_{\text{rel}} = t - t'$ is the relative time of the two time arguments in the Green's functions. The particle and hole occupation are

$$N(\omega) = A(\omega)f(\omega), \quad \bar{N}(\omega) = A(\omega)\bar{f}(\omega), \qquad \text{(S18)}$$

where $f(\omega) = \frac{1}{1+e^{\beta\hbar\omega}}$ is the Fermi distribution function and $\bar{f}(\omega) = 1 - f(\omega)$ for inverse temperature β.

In the nonequilibrium case, we measure the local spectral function and the occupation as

$$A(\omega;t_{\text{av}}) \equiv -\frac{1}{2\pi} \sum_{\alpha=A,B} \text{Im} \int dt_{\text{rel}} G^R_{ii,\alpha\alpha}(t_{\text{rel}};t_{\text{av}})e^{i\omega t_{\text{rel}}}, \qquad \text{(S19)}$$

$$N(\omega;t_{\text{av}}) \equiv \frac{1}{4\pi} \sum_{\alpha=A,B} \text{Im} \int dt_{\text{rel}} G^<_{ii,\alpha\alpha}(t_{\text{rel}};t_{\text{av}})e^{i\omega t_{\text{rel}}}. \qquad \text{(S20)}$$

Here $<$ indicates the lesser part of the Green's function, $t_{\text{av}} = \frac{t+t'}{2}$ is the average time of the two time arguments in the Green's functions and $G(t_{\text{rel}};t_{\text{av}}) = G(t,t')$. The hole occupation is defined as $\bar{N}(\omega;t_{\text{av}}) = A(\omega;t_{\text{av}}) - N(\omega;t_{\text{av}})$.

We also introduce the current operator along the x direction and its current-current response function,

$$\hat{J}_x = qJ_{||} \sum_{\mathbf{k},\sigma} \left[\hat{c}^\dagger_{\mathbf{k},A,\sigma}(t') \;\; \hat{c}^\dagger_{\mathbf{k},B,\sigma} \right] \hat{\sigma}_2 \begin{bmatrix} \hat{c}_{\mathbf{k},A,\sigma} \\ \hat{c}_{\mathbf{k},B,\sigma} \end{bmatrix}, \qquad \text{(S21)}$$

$$\chi_{xx}(t,t') = -i\langle T_{\mathcal{C}}\hat{J}_x(t)\hat{J}_x(t')\rangle. \qquad \text{(S22)}$$

Here $\hat{\sigma}_2$ is the Pauli matrix. The retarded part of $\chi_{xx}(t,t')$ is the response function, and the imaginary part of $\chi^R_{xx}(\omega)$ corresponds to the probability of energy absorption from weak excitation fields. If we neglect the vertex correction (this is exact in simple lattices such as cubic and hyper-cubic lattices), $\chi_{xx}(t,t')$ can be expressed with the single particle Green's functions as

$$\chi_{xx}(t,t') = -iq^2 J_{||}^2 \sum_{\mathbf{k}} \text{tr}[\hat{\sigma}_2\hat{G}_{\mathbf{k}}(t,t')\hat{\sigma}_2\hat{G}_{\mathbf{k}}(t',t)]. \quad \text{(S23)}$$

This leads to

$$-\frac{1}{\pi}\text{Im}\chi_{xx}(\Omega) = q^2 J_{||}^2 \sum_{\mathbf{k}} \int d\omega \text{tr}[\hat{\sigma}_2\hat{A}_{\mathbf{k}}(\omega+\Omega)\hat{\sigma}_2\hat{A}_{\mathbf{k}}(\omega)]$$
$$\times [f(\omega)\bar{f}(\omega+\Omega) - \bar{f}(\omega)f(\omega+\Omega)]. \qquad \text{(S24)}$$

Here $\hat{A}_{\mathbf{k}}(\omega) \equiv -\frac{1}{2\pi}\text{Im}[\hat{G}^R_{\mathbf{k}}(\omega) - \hat{G}^A_{\mathbf{k}}(\omega)]$, which is the single particle spectrum, and R and A indicate the retarded and advanced parts. The first term corresponds to absorption and the second term to emission. In the present study, due to $\hbar\Omega\beta \gg 1$ we can neglect the emission term. In addition, in the strong coupling regime, the support of $A_{\mathbf{k}}(\omega)$ turns out to be almost the same as $A(\omega)$, so that it is enough to consider $A(\omega)$ to roughly estimate the possible absorption process.

II. OFF-RESONANT DRIVING

In this part, we choose $J_x = 1, J_y = J_z = 0.2$ and $U = 3.75$, which correspond to 200 Hz, 40 Hz and 750 Hz, respectively. We use the inverse temperature of $\beta = 0.83$, which is estimated by the experiment in a preparation lattice (IV I). In addition to the initial lattice configuration stated above, $\mathcal{D}$ is also measured for a static reference lattice with a reduced tunneling $J_x = 0.4$ (corresponds to 80 Hz) in the experiment. Since the interaction is the same for both configurations, U/W increases from 1.1 for $J_x = 1$ to 1.6 for $J_x = 0.4$. With the larger value of U/W, the temperature estimation at equal entropy is more difficult, hence, direct evaluation of the reference value from DMFT becomes also difficult. Thus, the best estimation for $\mathcal{D}$ in local equilibrium is provided by an adiabatic ramp from $J_x = 1$ to $J_x = 0.4$. We use a ramp time of 5 ms where saturation at $\mathcal{D} = 0.043$ is reached (see Fig. 2).

In the lowest order $1/\Omega$-expansion, the effective static Hamiltonian (after the ramp) is obtained by neglecting the oscillating terms in Eq. (S6), where the hopping parameter of particles along the x direction is renormalized by $\mathcal{J}_0(K_0)$. The bandwidth of the free particle system is thus reduced from $W = 2J_{||} + 4J_{\perp}$ to $W' = 2\mathcal{J}_0(K_0)J_{||} + 4J_{\perp}$. The neglected terms can be regarded as additional excitation terms of this effective static Hamiltonian.

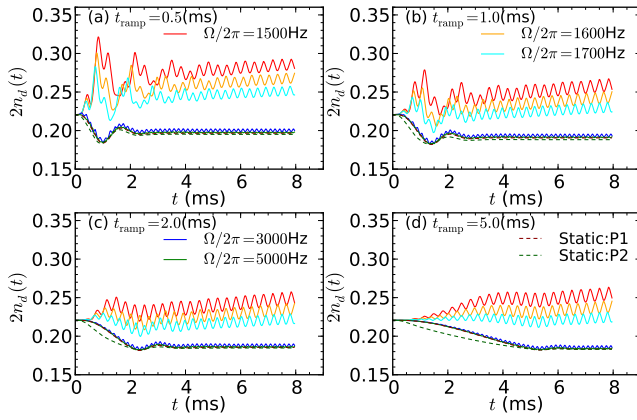In the experiment, the external field is introduced with finite ramp times, Eq. (S5). If the ramp speed is

FIG. S3. Time evolution of the double occupation $2n_d(t)$ for different excitation protocols and ramp times. Solid lines correspond to the time-periodic excitations with specified frequencies, while the dashed lines correspond to the modulations of the hopping parameter (see the text). Here, $\beta = 0.83$, $K_0 = 1.43$ and $n = 1.0$. P1 and P2 stands for protocol 1 and protocol 2.

slow enough, it is expected that the dynamics can still be described by an effective Hamiltonian with a time-dependent hopping in the x direction but without time-periodic excitation,

$$
- J_{||}\mathcal{J}_0(K_0) \sum_{\langle \mathbf{i},\mathbf{j}\rangle_x,\sigma} \hat{c}^\dagger_{\mathbf{i},\sigma}\hat{c}_{\mathbf{j},\sigma}
$$
$$
\to - J_{||}\mathcal{J}_0(K_0 \cdot F_{\mathrm{ramp}}(t;t_{\mathrm{ramp}})) \sum_{\langle \mathbf{i},\mathbf{j}\rangle_x,\sigma} \hat{c}^\dagger_{\mathbf{i},\sigma}\hat{c}_{\mathbf{j},\sigma}. \quad (S25)
$$

We call this the "protocol 1" excitation. On the other hand, in the experiments, a linear ramp of the hopping parameter in the x direction is implemented:

$$
- J_{||}\mathcal{J}_0(K_0) \sum_{\langle \mathbf{i},\mathbf{j}\rangle_x,\sigma} \hat{c}^\dagger_{\mathbf{i},\sigma}\hat{c}_{\mathbf{j},\sigma} \quad (S26)
$$
$$
\to - J_{||}(1 + (\mathcal{J}_0(K_0) - 1) \cdot F_{\mathrm{ramp}}(t;t_{\mathrm{ramp}})) \sum_{\langle \mathbf{i},\mathbf{j}\rangle_x,\sigma} \hat{c}^\dagger_{\mathbf{i},\sigma}\hat{c}_{\mathbf{j},\sigma}.
$$

We call this the "protocol 2" excitation. In the following, we compare the results of the periodic excitation and these two excitation protocols.

A. Real time propagation

In this section, we present the time-evolution of the double occupation $(2n_d(t))$ simulated by DMFT with a fixed density. In Fig. S3, we show the results for different excitation protocols and ramp times at half-filling. For all ramp times studied here, the results of $\Omega/(2\pi) = 3000$ Hz and $\Omega/(2\pi) = 5000$ Hz almost match those of the hopping modulation protocol 1, with $\Omega/(2\pi) = 5000$ Hz showing a better agreement, as expected. In particular, it is interesting to point out that the description based
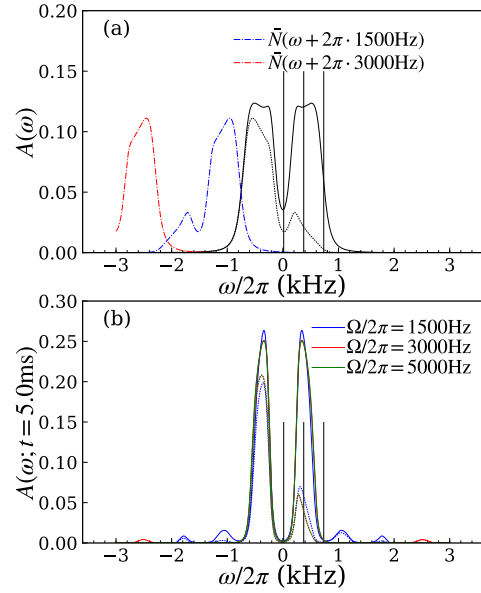

FIG. S4. (a) Local single-particle spectral function ($A(\omega)$, black curve) and its occupation ($N(\omega)$, black dashed curve) at $\beta = 0.83$ in equilibrium at half-filling. The unoccupied parts $\bar{N}$ of the spectral function shifted by 1500 Hz and 3000 Hz are also shown to illustrate possible absorption processes in the driven state (dash-dotted lines). (b) Nonequilibrium local single-particle spectral function measured at $t_{\mathrm{av}} = 5.0$ ms for $K_0 = 1.43$, $t_{\mathrm{ramp}} = 5.0$ ms, $n = 1$ and different excitation frequencies Ω. The vertical lines indicate $\frac{U}{2} - \frac{W}{2}$, $\frac{U}{2}$ and $\frac{U}{2} + \frac{W}{2}$.

on the static Floquet Hamiltonian, time-dependent effective hopping parameter is meaningful even when the amplitude of the excitation is modulated rather quickly compared to the excitation frequency. For example, for $t_{\mathrm{ramp}} = 1.0$ ms, there are only three oscillations for $\Omega/(2\pi) = 3000$ Hz and five oscillations for 5000 Hz during the ramp of the field strength. Nevertheless, the dynamics predicted by the Floquet Hamiltonian reproduces the full dynamics rather well. The results of the hopping modulation protocol 2 show a somewhat different dynamics during the ramp, but the values of the double occupation reached at the end of ramp and the dynamics after the ramp are very similar to the results obtained using the protocol 1. This tendency is particularly clear as we increase the ramp time.

Now we reduce the excitation frequency down to 1500 Hz, which is just above the naively expected bandwidth $(U + W)/h = 1470$ Hz and that of the Floquet Hamiltonian after the ramp $(U + W')/h = 1230$ Hz. We clearly see large deviations from the prediction of the Floquet Hamiltonian, and in particular a double occupation which continuously increases. There are several points to note here. Firstly, around this frequency range the nonequilibrium dynamics strongly depends on the excitation frequency, compare the results for $1500, 1600, 1700$ Hz. Secondly, the production of double

occupations is particularly pronounced during the ramp, while after the ramp the increase becomes more moderate. Thirdly, the shorter the ramp time, the larger the increase of the double occupation during the ramp. The last point manifests itself as a larger value of $2n_d$ averaged over one period after the ramp in the experiment. The first and second points can be related to the local spectral function, and we come back to this point in the next section. The third observation is explained by the fact that for the shorter ramp, the field includes higher frequency components than Ω during the ramp, which can lead to efficient absorption of energy from the field. We also note that the results away from half-filling ($n = 0.6$) show the same qualitative behavior, including the same characteristic time scales.

B. Analysis of the single-particle spectrum

Here, we discuss the properties of the single-particle spectrum of the many-body state, which is difficult to access experimentally. In Fig. S4(a), we show the DMFT result for the local spectrum ($A(\omega)$) and its occupation in equilibrium at half filling (black curves). Compared to the naive expectation that the total bandwidth is just $U + W$, it is larger because of correlation-induced broadening. In the figure, we also show the spectral functions shifted by 1500 Hz and 3000 Hz. As explained above, the overlap between the shifted unoccupied part of the spectrum and the occupation can be roughly connected to the efficiency of the absorption under a weak excitation with frequency corresponding to the shift. For 1500 Hz, there is a non-negligible overlap because of the broadened spectral function, while for 3000 Hz there is almost no overlap. These results are consistent with the doublon dynamics observed in the previous section. In particular, when the excitation frequency is comparable to the width of the spectrum, a small change of the frequency substantially changes the overlap area, which leads to a sensitive dependence of the dynamics on the excitation frequency.

Moreover, when the excitation amplitude becomes stronger, the bandwidth of the upper and lower bands is reduced, which leads to a reduction of the total bandwidth and hence the strength of the absorption. To illustrate this effect, we plot in Fig. S4(b) the local spectral function $A(\omega; t)$, measured after completion of the ramp, which indeed exhibits a substantial reduction of the bandwidth. This can explain the more moderate increase of the double occupation after the ramp.

III. NEAR-RESONANT DRIVING

Here we consider the case where $U \approx \hbar\Omega$, where the efficient creation of $\mathcal{D}$ is expected [S3, S4]. In this case, we can introduce the effective static Hamiltonian Eq. (3) in the rotating frame [S5]. In this section, we choose $J_x = 1, J_y = J_z = 0.5$ and $U = 17.5$, which correspond
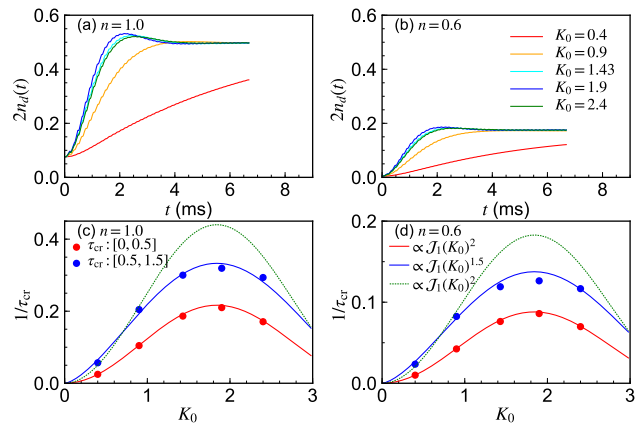

FIG. S5. (a)(b) Time evolution of the double occupation $2n_d(t)$ after the AC quench with different field strengths. (c)(d) Doublon creation rates estimated from the slopes of the linear fits for the indicated time intervals and for different field strengths. We also show lines proportional to $\mathcal{J}_1(K_0)^2$ and $\mathcal{J}_1(K_0)^{1.5}$. Panel (a)(c) is for $n = 1.0$ and panel (b)(d) is for $n = 0.6$. Here, $U/h = \Omega/(2\pi) = 3500$ Hz and $\beta = 0.30$.

to 200 Hz, 100 Hz and 3500 Hz, respectively. We use the inverse temperature of $\beta = 0.30$, which is estimated by the experiment (IV I).

A. Quench dynamics: Doublon creation ratio

Before we show the time evolution under the excitation with finite ramp time, we analyze the dynamics after an AC quench for $U = \hbar\Omega$, i.e. $t_{\text{ramp}} = 0$. In this case, the effective Hamiltonian (in the rotating frame) exhibits no time dependence, and it implies that the hopping parameter relevant for the doublon-holon creation is proportional to $\mathcal{J}_1(K_0)$. One can reach the same conclusion from Eq. (S6) by regarding the $n = \pm 1$ terms as perturbations and neglecting higher harmonic terms. Since $\mathcal{J}_1(K_0)$ exhibits a maximum around $K_0 \simeq 1.85$, the doublon creation rate should exhibit a maximum around this value. Indeed, we can see this in the DMFT analysis for different fillings, see Fig. S5(a)(b). The fastest increase of the doublon number is observed for $K_0 = 1.9$ regardless of the density. Initially, the increase shows a super-linear (quadratic) behavior for all cases considered here. This behavior can be understood by considering the time evolution from the state $\hat{c}_{1,\uparrow}^\dagger \hat{c}_{2,\downarrow}^\dagger |\text{vac}\rangle$ in a dimer system. The time evolution in this system results in an increase of the doubly occupied states $\propto \sin^2(2J_x\mathcal{J}_1(K_0)t/\hbar)$.

We also extracted the doublon creation rate by linear fitting, $n_d(t) = t/\tau_{\text{cr}} + a$. In Fig. S5(c)(d), we show $1/\tau_{\text{cr}}$ extracted using two different fitting ranges. One can see that at the very early stage, $1/\tau_{\text{cr}}$ scales as $\mathcal{J}_1^2(K_0)$, as is expected from the Femi golden rule. On the other hand, $1/\tau_{\text{cr}}$ deviates from $\mathcal{J}_1^2(K_0)$ and better matches $\mathcal{J}_1^{1.5}(K_0)$ at later times.
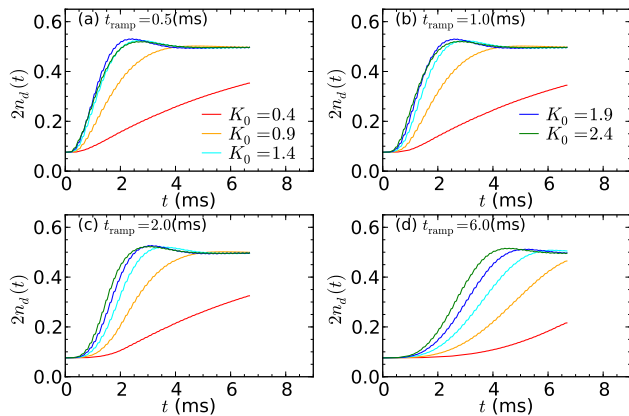
FIG. S6. Time evolution of the double occupation $2n_d(t)$ for different field strengths and ramp times. Here, $U/h = \Omega/(2\pi) = 3500$ Hz, $\beta = 0.30$ and $n = 1.0$.

B. Real-time propagation

Now we move on to the results for the finite ramp time, which are relevant for the comparison to the experiments. In Fig. S6, we show the time evolution of the double occupation $2n_d(t)$ for different field strengths and ramp times at half filling. When $t_{\rm ramp}$ is small, one can see the fastest increase at $K_0 = 1.9$ as in the quench case, see $t_{\rm ramp} = 0.5$ ms. As the ramp time increases, $K_0 = 2.4$ starts to show a faster increase. This is because for larger K_0 the system experiences a "decent field strength" for a longer period of time during the ramp. In practice, $K_0 \in [1.4, 2.4]$ results in a fast increase of similar magnitude. If we compare the ramp of the field up to $K_0 = 1.9$ to the ramp up to $K_0 = 2.4$ for the same ramp time, the latter includes a longer time period where K_0 is in $[1.4, 2.4]$. The results for the doped system ($n = 0.6$) show the qualitatively same behavior, including the same characteristic time scales.

In Fig. S7, we show the effects of detuning around $U \simeq \hbar\Omega$. The fastest increase of the double occupation can be observed at $U = \hbar\Omega$. At half-filling, the sign of the detuning $U - \hbar\Omega$ has some influence on the doublon production, but the difference between the $U - \hbar\Omega > 0$ and $U - \hbar\Omega < 0$ regimes is not so large. We point out that a similar dependence of the doublon creation was observed near the other resonant condition $U \simeq 2\hbar\Omega$ in [S3], where the difference between the two regimes was shown to becomes negligible as U increases. On the other hand, away from half-filling, the asymmetry between the positive and negative detuning results becomes prominent, while the difference between $U/h = 3000$ Hz and $U/h = 3500$ Hz becomes less prominent, which can be explained by analyzing the properties of the spectral functions, as discussed in the next section. We also note that the difference between positive and negative detuning becomes more prominent for lower temperatures.

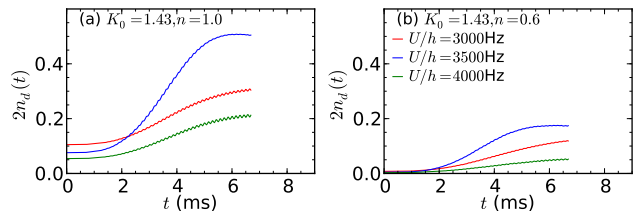In Fig. S8, we show the time evolution of the double

FIG. S7. Time evolution of the double occupation $2n_d(t)$ for different U around $\Omega/(2\pi) = 3500$ Hz. Here, we choose $t_{\rm ramp} = 6.0$ ms, $K_0 = 1.43$, and $\beta = 0.30$.
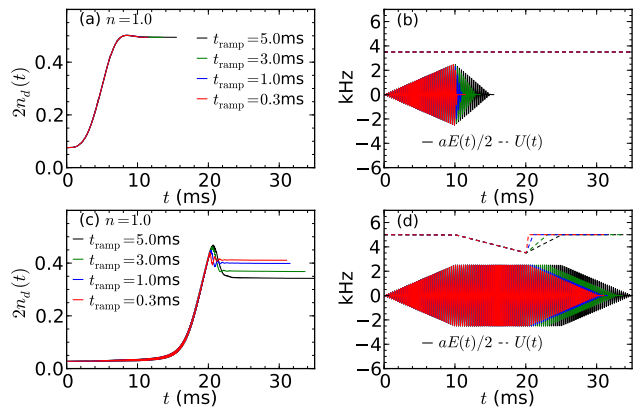

FIG. S8. Time evolution for different adiabatic protocols at half-filling. (a) Double occupation $2n_d(t)$ for the simple adiabatic protocols with different ramp-down times ($t_{\rm ramp}$), whose field strength and interaction strength are shown in panel (b). (c) Double occupation $2n_d(t)$ for the sophisticated protocols with different ramp-down times ($t_{\rm ramp}$), whose field strength and time-dependent interaction strength are shown in panel (d). Here, we choose $K_0 = 1.43$, and $\beta = 0.30$.

occupation for different excitation protocols to check the adiabaticity. Figure S8(a) depicts the evolution under the simple protocol, see also Fig. S13 below. One can see that the final value of the double occupation is almost independent of the ramp-down time. On the other hand, under the sophisticated protocol (panel (c)(d)), where the driving is ramped up away from resonance and the interaction is subsequently ramped into resonance (and vice versa for the ramp down), one can clearly see a reduction of the final value of the double occupation as the ramp-down time of the interaction is increased. These results are consistent with the experimental observation, where the sophisticated method shows a rather prominent reduction of the double occupation with increasing ramp-down time, see Fig. S14.

C. Analysis of the single-particle spectrum

In Fig. S9, we show the local single-particle spectral functions and corresponding occupations for different fillings at $U/h = 3500$ Hz. At half filling, one can ob-
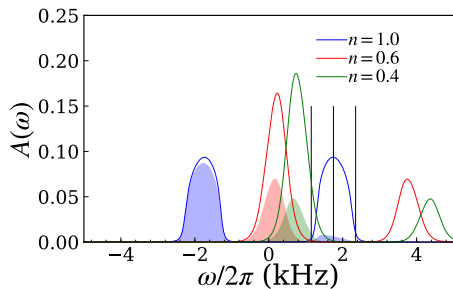
FIG. S9. Local single-particle spectral function ($A(\omega)$) and its occupation ($N(\omega)$, shaded area) at $\beta = 0.3$ in equilibrium for different fillings. The vertical lines indicate $\frac{U}{2} - \frac{W}{2}$, $\frac{U}{2}$ and $\frac{U}{2} + \frac{W}{2}$.

serve clear Hubbard bands, whose centers are located near $\omega = \pm \frac{U}{2\hbar}$ and the bandwidth of each band is essentially W. We note that the band edges become sharper as temperature is decreased.

Away from half-filling, with finite hole doping, the position of the lower band is located around $\omega = 0$ and the lower part of the lower Hubbard band is occupied. Still the distance between the bands is almost U and their bandwidth is similar. The weight of the upper band is decreased from $\frac{1}{2}$, while that of the lower band is increased from $\frac{1}{2}$. This is because the weight of the upper band, roughly speaking, reflects the probability of producing a doublon by adding a particle to the system. This probability is reduced if the system is hole doped.

Now we show that the effect of detuning mentioned above can be well understood by considering the structure of the spectral functions and occupations. As can be seen in Eq. (S24), the absorption ratio is associated with $\mathrm{tr}[\hat{\sigma}_2 \hat{A}_k(\omega + \Omega) \hat{\sigma}_2 \hat{A}_k(\omega)]$. In the present case, we numerically confirmed that the contribution from the off-diagonal component of $\hat{A}$ is much smaller than that from the diagonal terms. Therefore, one can focus on $A_{\mathbf{k},AA}(\omega) = A_{\mathbf{k},BB}(\omega)$. These spectral functions consist of two bands, which are separated by U and have almost the same bandwidth as in $A(\omega)$. In particular, at half-filling, the shapes of the upper band and the lower band in $A_{\mathbf{k}}(\omega)$ (and hence in $A(\omega)$) turn out to be almost identical. Because of this, at half-filling, the lower band (occupied band) and the upper band (unoccupied band) overlap almost perfectly when they are shifted by $\hbar\Omega = U$ and this frequency shows the largest value of $|\mathrm{Im}\chi_{xx}(\Omega)|$. Positive and negative detuning can reduce the overlap. However, because of the similar shapes of the upper and lower bands, the overlap is similar and the value of $\mathrm{Im}\chi_{xx}(\Omega)$ also becomes similar for positive and negative detuning. In the hole-doped case, the lower part of the lower band is occupied, and these states can be excited to the lower part of the upper band when $\hbar\Omega = U$. For negative detuning $\hbar\Omega > U$, there exist final states in the upper band. On the other hand, for positive detuning, the number of accessible final states
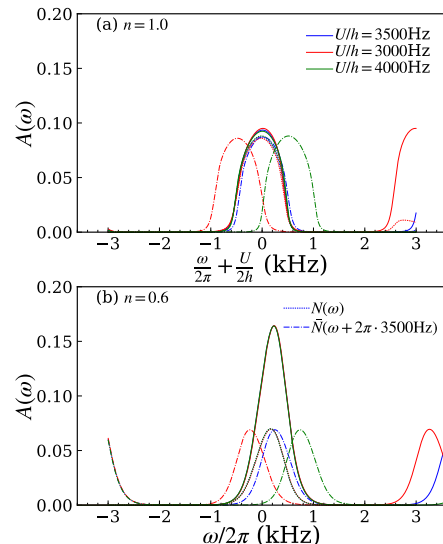

FIG. S10. Local single-particle spectral function ($A(\omega)$), its occupation ($N(\omega)$, dashed lines) and the spectral functions shifted by indicated frequencies (dash-dotted lines) at $\beta = 0.30$ in equilibrium. Panel (a) is for $n = 1.0$ and panel (b) is for $n = 0.6$. The overlap between the occupation and the shifted spectrum for the unoccupied states roughly shows the efficiency of absorption under the excitation with the corresponding frequency.

is reduced because the occupation shifted by Ω may be located within the gap.

We can directly demonstrate the above point using the local spectrum as a representative of $A_{\mathbf{k}}$. In Fig. S10 (a), we show the overlap between the occupation and the unoccupied states shifted by $\hbar\Omega$ for different U at half filling. One can see that the overlap is maximal at $U = \hbar\Omega$, and that for $U/h = 3000$ Hz and $U/h = 4000$ Hz it is almost the same. Away from half-filling, one can identify clear differences in these overlaps between $U/h = 3000$ Hz and $U/h = 4000$ Hz. This originates from the larger occupation in the lower part of the lower band and the distance of the bands being almost U. These observations naturally explain the effects of detuning on the speed of the doublon creation observed in Fig. S7.

IV. EXPERIMENT

A. General preparation

We start with a gas of ^{40}K fermions in the $m_F = -9/2, -5/2$ sublevels of the $F = 9/2$ manifold confined in a harmonic optical dipole trap. The mixture is evaporatively cooled down to a spin-balanced Fermi degenerate gas of $35(3) \times 10^3$ atoms at a temperature $T/T_F = 0.14(2)$ (T_F is the Fermi temperature). Feshbach resonances (around 202.1 G and 224.2 G) allow us to tune the interactions from weakly to strongly repulsive using a $-9/2, -7/2$ or a $-9/2, -5/2$ mixture, respectively. The

latter is prepared with a Landau-Zener transfer across the $-7/2 \rightarrow -5/2$ spin resonance.

Four retro-reflected laser beams of wavelength $\lambda = 1064$ nm are used to create the optical lattice potential

$$
\begin{aligned}
V(x,y,z) = &-V_{\overline{X}}\cos^2(kx + \theta/2) - V_X\cos^2(kx) \\
&-V_{\tilde{Y}}\cos^2(ky) - V_Z\cos^2(kz) \\
&-2\alpha\sqrt{V_X V_Z}\cos(kx)\cos(kz)\cos\varphi, \text{(S27)}
\end{aligned}
$$

where $k = 2\pi/\lambda$ and x, y, z are the laboratory frame axes. The lattice depths $V_{\overline{X},X,\tilde{Y},Z}$ (in units of the recoil energy $E_R = h^2/2m\lambda^2$, h is the Planck constant and m the mass of the atoms) are independently calibrated using amplitude modulation on a ^{87}Rb Bose-Einstein condensate (BEC). The visibility $\alpha = 0.99(1)$ is calibrated using the same technique but in an interfering lattice configuration with $V_X, V_Z \neq 0$. The phases θ and φ are set to $1.000(2)\pi$ and $0.00(3)\pi$, respectively and determine the lattice geometry. The experimental Hubbard parameters J and U are numerically calculated from the lattice potential's Wannier functions computed with band-projected position operators [S6]. The bandwidth W is defined as $W = 2(J_x + 2(J_y + J_z))$ for the single band tight-binding model. Tunneling terms J_x, J_z are along the horizontal and vertical links of the hexagonal unit cell, and J_y along the orthogonal direction of the hexagonal plane, see Fig. S15. The tunneling link J_w across the hexagonal unit cell is strongly suppressed. We note that the tunneling parameters $J_\parallel, J_\perp$ introduced in the theoretical studies are related to the ones presented here for the experiments by $J_\parallel = J_x$ and $J_\perp = J_y = J_z$. The estimated values for the Hubbard parameters used in the experiments, namely tunnelings $J_{x,y,z}$, bandwidth W and on-site interaction U, are given in Tables I to IV for the different driving protocols used in the experiment, which were also used for the theoretical simulations.

B. Periodic driving

The driving protocols are carried out with a piezoelectric actuator that periodically modulates in time the position of the X retro-mirror at a frequency $\Omega/(2\pi)$ and amplitude A. As a result, the retro-reflected X and $\overline{X}$ laser beams acquire a time-dependent phase shift with respect to the incoming beams and the lattice potential is $V(x - A\cos(\Omega t), y, z)$. The normalized amplitude is given by $K_0 = mA\Omega d_x/\hbar$, with d_x the inter-site distance along the x axis ($\hbar = h/2\pi$). Values of K_0 are given in tables I to IV for all the driving protocols presented in this work. These include a factor of d_x corresponding to the distance between Wannier functions located on the left and right sites of a lattice bond for the corresponding lattice geometry.

In addition, the phases of the incoming X and Z lattice beams are modulated at frequency $\Omega/(2\pi)$ using acousto-optical modulators to stabilize the phase $\varphi = 0.00(3)\pi$. This compensation is not perfect and leads to a residual

amplitude modulation of the lattice beams. Also, the phase and amplitude of the mirror displacement driven by the piezoelectric actuator is calibrated by measuring the amplitude of the phase modulation without any compensation.

C. Detection methods

The detection of double occupancies starts with freezing the dynamics by ramping up the lattice depths to $V_{\overline{X},X,\tilde{Y},Z} = [30, 0, 40, 30]$ E_R within 100 μs. Freezing the lattice at different modulation times within one driving period allows us to average over the micromotion for each measurement protocol. Later, we linearly ramp off the periodic driving within 10 ms. Then, a (radio-frequency) Landau-Zener transition tuned to an interaction-dependent spin resonance selectively flips the $m_F = -7/2$ atoms on doubly occupied sites to the initially unpopulated $m_F = -5/2$ spin state (or $m_F = -5/2$ is flipped into the $m_F = -7/2$ state when we start with a $-9/2, -5/2$ mixture). Subsequently, we perform a Stern-Gerlach type measurement after switching off all confining potentials to separate the individual spin states. After 8 ms of ballistic expansion the $-9/2, -7/2, -5/2$ spin components are spatially resolved in an absorption image. For each spin component its spatial density distribution is fitted with a Gaussian profile to estimate the fraction of atoms on each spin state determining the fraction of double occupancies $\mathcal{D}$.

D. Off-resonant modulation

Once the degenerate Fermi gas is prepared, the optical lattice is loaded by ramping up the lattice beams within 200 ms to an anisotropic configuration $J_{x,y,\tilde{z}}/h = [487, 102, 83]$ Hz (see table V and Fig. S15 for definition of $J_{\tilde{z}} = \tilde{J}_z = (2J_z + J_w)/3$). From this intermediate loading configuration we then ramp within 10 ms to a hexagonal configuration with reduced tunnelings $J_{x,y,z}/h = [191(31), 41(3), 40(3)]$ Hz (see table I). The chosen values for the tunneling enables us to resolve dynamics on the order of the driving period. We leave the harmonic trapping potential and the anisotropy constant during the second loading process such that we get a locally equilibrated state. During this loading the interactions in a $m_F = -9/2, -7/2$ mixture are also set, see *starting lattice* in Table I. In this configuration the system is prepared at $U/h = 753(26)$ Hz and $W/h = 707(65)$ Hz, right in the crossover regime from a metal to a Mott insulator. Then, the off-resonant driving protocol starts by ramping up K_0 from 0 to $1.70(2)$ within a variable amount of time t_ramp. The modulation is held very shortly (100 μs) before the detection of $\mathcal{D}$ starts. At ramp times below 100 μs (shaded area in Fig. 2) residual dynamics during the detection process and the finite bandwidth of the piezoelectric actuator influence

Off-resonant modulation

parameter	starting lattice	final lattice
Driven system		
$J_{x,y,z}/h$ (Hz)	191(31), 41(3), 40(3)	-
W/h (Hz)	707(65)	-
$J_{x,y,z}^{\mathrm{eff}}/h$ (Hz)	-	76(13), 41(3), 40(3)
W^{eff}/h (Hz)	-	478(32)
U/h (Hz)	753(26)	
K_0	1.70(2)	
Undriven system		
$J_{x,y,z}/h$ (Hz)	193(34), 41(3), 39(4)	81(13), 41(3), 39(4)
W/h (Hz)	708(71)	481(33)
U/h (Hz)	752(27)	
K_0	-	

TABLE I. Experimental parameters for the off-resonant modulation protocols in the driven and undriven cases. Errors in $(J_{x,y,z},\ W)$ and K_0 account for the uncertainty of the measured lattice depths $V_{\overline{X},X,\tilde{Y},Z}$ and uncertainty in the estimated lattice bond distance d_x, respectively. The error in U additionally includes the uncertainty on the measured magnetic field and Feshbach resonance calibration. $J_{x,y,z}^{\mathrm{eff}}$ (and so W^{eff}) are computed from the relation $J^{\mathrm{eff}} = J\mathcal{J}_0(K_0)$.
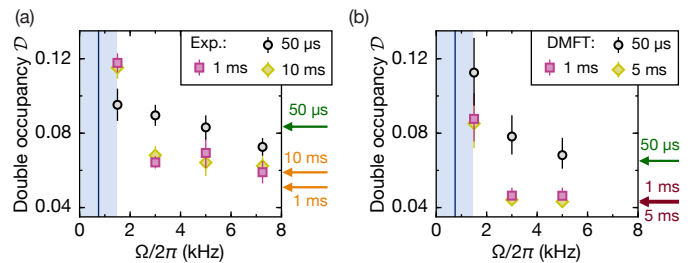

FIG. S11. Crossover from off-resonant to near-resonant modulation. Experimental (a) and DMFT (b) data display $\mathcal{D}$ measured after ramping up the drive for different frequencies. The vertical line indicates the interaction $U/h = 0.75(3)$ kHz and the blue area around it has the width of twice the bandwidth $W/h = 0.71(7)$ kHz. The data is presented for three different ramp times corresponding to 50 μs (circles), 1 ms (squares) and 10 ms (5 ms) (diamonds) for panel a (b). Arrows labelled with holding time indicate data taken in a static system. The upper green arrow shows the reference value holding in the initial ($J_x/h = 193(34)$Hz) lattice for 50 μs. In the experimental data the two lower orange arrows denote the reference value for the final lattice ($J_x/h = 81(13)$ Hz) after holding 1 ms (lower value) and 10 ms (upper value). The lower red arrows in the DMFT data represent the value after a 1 ms and 5 ms ramp of the tunneling from initial to final. The $\mathcal{D}$ saturates at 1 ms and only a negligible change compared to 5 ms is seen. Error bars in (a) denote the standard error for 5 measurements and in (b) reflect the uncertainty of the entropy estimation in the experiment [39].

the measured value of $\mathcal{D}$. The level of $\mathcal{D}$ is compared to the one of an undriven (static counterpart) protocol. This static protocol consists in ramping down (up) the V_X (V_Z) lattice depth (*final lattice* configuration), which results in a ramp of J_x/h from 193(34) Hz to 81(13) Hz. We note that this tunneling ramp does not exactly follow the 0th-order Bessel function tunneling renormalization $J^{\mathrm{eff}}[t_{\mathrm{ramp}}] \sim \mathcal{J}_0\left(K_0[t_{\mathrm{ramp}}]\right)$ expected in the driven protocol. Further theoretical studies presented above (section III A) show that the differences in the dynamics are negligible between these two protocols. Two reference levels for $\mathcal{D}$ are additionally measured when we simply load the atoms in a static lattice of either $J_x/h = 193(34)$ Hz (*starting lattice*) or $J_x/h = 81(13)$ Hz (*final lattice*).

In Fig. 2 of the main text it is shown how the effective description breaks down when the drive frequency $\Omega/2\pi$ approaches the on-site interaction U. We present the same data in Fig. S11 showing $\mathcal{D}$ as a function of drive frequency. The data displayed is taken after a ramp time of 50 μs (circles), 1 ms (squares) and 10 ms for the experimental part (a) or 5 ms for the DMFT calculations (b) (diamonds). The solid vertical line denotes the value of the interaction and the blue area around it has the width of twice the bandwidth. In this region, which corresponds to the near-resonant driving regime, direct excitations of $\mathcal{D}$ are possible. The upper green arrow denotes the reference value taken after holding 50 μs in the initial static lattice. The two lower orange arrows in panel (a) depict reference values in a static lattice with a tunneling of

$J_x/h = 81(13)$ Hz after holding for 1 ms (lower) or 10 ms (upper). In panel (b) the lower red arrows correspond to the $\mathcal{D}$ reached after a ramp of the tunneling from the intial to the final value in 1 ms and 5 ms. The difference of these two values is negligible indicating the saturation in $\mathcal{D}$ after a 1 ms ramp.

The effective Hamiltonian predicts a change of $\mathcal{D}$ from the reference value in the initial lattice (arrow denoted with 50 μs) to the reference value of the final lattice (arrows denoted with longer hold times). For frequencies of 3 kHz and higher the measured and calculated $\mathcal{D}$ are within the expectation of the effective Hamiltonian whereas for 1.5 kHz a much larger value of $\mathcal{D}$ appears. We attribute the remaining deviations in the values of $\mathcal{D}$ between the ab-initio calculations and the experimental values to the systematic uncertainties on the input temperature and density profiles provided by the experiment. Experimental data taken at 50 μs is in addition influenced by residual dynamics during the detection and the finite bandwidth of the piezoelectric actuator.

E. Adiabaticity in off-resonant modulation

We perform tests on the adiabaticity to revert the suppression of $\mathcal{D}$ in the off-resonant protocols, see Fig. S12. Here, the system is prepared in a driven lattice ($\Omega/(2\pi) = 7.25$ kHz) where we use a 10 ms ramp up of K_0 to a
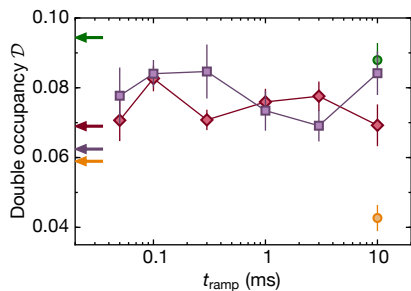
FIG. S12. Adiabaticity in off-resonant modulation. In the driven system with $\Omega/(2\pi) = 7.25$ kHz (purple squares) K_0 is ramped up in 10 ms, held for 100 μs and then ramped down within a time t_{ramp}. In the undriven system (red diamonds) the lattice depths are directly ramped within the same time as for the driven case. Reference levels of $\mathcal{D}$ with $t_{\text{ramp}} = 0$ s are shown (arrows). For comparison, the level of $\mathcal{D}$ in the static systems with $J_x/h = 79(12)$ Hz (yellow circle) and $J_x/h = 208(34)$ (green circle) is shown when the waiting times are [10 ms + 100 μs] (arrows) or [10 ms + 100 μs + 10 ms] (circles). Data points are the mean and standard error of 5 individual runs (including an average over one driving period $(2\pi)/\Omega$ in the data versus time).

final value of 1.68(2). From this *starting lattice* configuration with $U/h = 752(26)$ Hz and $J_{x,y,z}^{\text{eff}}/h = [85(15), 42(4), 38(4)]$ Hz (see Table II) we hold 100 μs before K_0 is ramped down to 0 within a variable amount of time t_{ramp} reaching the *final lattice* configuration where $J_x/h = 208(37)$ Hz. The detected level of $\mathcal{D}$ is compared to an undriven protocol that ramps the lattice depths V_X and V_Z to emulate the tunneling renormalization in the driven protocol. This changes $J_{x,y,z}/h = [81(12), 42(3), 39(3)]$ Hz in the *starting lattice* to $J_x/h = 193(32)$ Hz in the *final lattice* configuration. We observe that $\mathcal{D}$ reaches comparable values in the driven and undriven systems and we do not see a clear change versus ramp time. Reference levels of $\mathcal{D}$ were also measured when the system is prepared in a static lattice with $J_x/h = 79(12)$ Hz or $J_x/h = 208(34)$ Hz. For each of these measurements the level of $\mathcal{D}$ is detected [10 ms + 100 μs] or [10 ms + 100 μs + 10 ms] after the lattice loading is done. These waiting times simulate the time that it takes to ramp up (10 ms), hold (100 μs) and ramp down (10 ms) the driving amplitude K_0. We attribute the remaining dynamics to the adiabatically imperfect preparation.

F. Near-resonant modulation

In the near-resonant driving protocols used to study the creation of $\mathcal{D}$ versus K_0 and U we load the cloud in 200 ms in an anisotropic lattice with $J_{x,y,\tilde{z}}/h = [202, 100, 67]$ Hz (see table V, and Fig. S15 for definition of $J_{\tilde{z}} = \tilde{J}_z = (2J_z + J_w)/3$). We further ramp in 10 ms to our starting hexagonal lattice configuration

Adiabaticity in off-resonant modulation

parameter	starting lattice	final lattice
	Driven system	
$J_{x,y,z}^{\text{eff}}/h$ (Hz)	85(15), 42(4), 38(4)	-
W/h (Hz)	491(36)	-
$J_{x,y,z}/h$ (Hz)	-	208(37), 42(4), 38(4)
W^{eff}/h (Hz)	-	737(76)
U/h (Hz)	752(26)	
K_0	1.68(2)	
	Undriven system	
$J_{x,y,z}/h$ (Hz)	81(12), 42(3), 39(3)	193(32), 42(3), 39(3)
W/h (Hz)	485(33)	711(71)
U/h (Hz)	750(27)	
K_0	-	

TABLE II. Experimental parameters for the adiabaticity in off-resonant modulation protocols of the driven and undriven cases. Note that the parameters for the *starting lattice* (*final lattice*) configuration are the ones of the *final lattice* (*starting lattice*) configuration of Table I, because for the adiabaticity protocols we explore the ramping down of K_0. Each parameter value and error was computed as described in Table I.

$J_{x,y,z}/h = [209(38), 100(7), 98(7)]$ Hz. In contrast to the off-resonant case we do not need to scale down the energy scales which results in a better preparation with respect to thermalization of the initial state. To reach strong repulsive interactions we use a $m_F = -9/2, -5/2$ mixture with its Feshbach resonance situated at 224.2 G. In particular we use $U/h = 3.5(1)$ kHz to prepare the system in a deep Mott insulating state. The driving amplitude K_0 is then ramped up during a varying time and the modulation is held for 100 μs at constant K_0 before the detection of $\mathcal{D}$ starts. Here $\Omega/(2\pi) = 3.5$ kHz in all protocols. Measurements with different K_0 are taken with fixed static Hubbard parameters, see Table III. Measurements with different U are taken with fixed $K_0 = 1.44(2)$.

G. Adiabaticity in near-resonant modulation

For the near-resonant driving protocols the adiabaticity to revert the creation of $\mathcal{D}$ is tested in two different measurement setups. In a first set of measurements the ramp up time of K_0 is fixed to 10 ms, the driving is held for another 100 μs at constant $K_0 = 1.44(2)$, and then K_0 is ramped down to 0. This is done for a varying K_0 ramp down time t_{ramp}, and the measurement is repeated at different interactions, see Table IV. Reference levels of $\mathcal{D}$ are taken right after the ramp up of K_0, i.e. before the ramp down, to compare the amount of $\mathcal{D}$ that is adiabatically reverted. We observe that $\mathcal{D}$ does not revert to the initial value for all ramp times t_{ramp}, i.e.,

Near-resonant modulation

parameter	value(s)
$\mathcal{D}$ versus K_0 protocol	
$J_{x,y,z}/h$ (Hz)	209(38), 100(7), 98(7)
W/h (Hz)	1210(85)
U/h (kHz)	3.5(1)
K_0	0, 0.40(1), 0.90(1), 1.43(2), 1.90(2), 2.40(3)
$\mathcal{D}$ versus U protocol	
$J_{x,y,z}/h$ (Hz)	195(35), 100(7), 102(8)
W/h (Hz)	1198(81)
U/h (Hz)	2.5(1), 3.0(1), 3.5(1), 4.0(1), 4.5(1)
K_0	1.44(2)

TABLE III. Experimental parameters for the near-resonant modulation protocols to measure $\mathcal{D}$ versus either K_0 or U^{eff}. The parameter values and errors were computed as described in Table I.
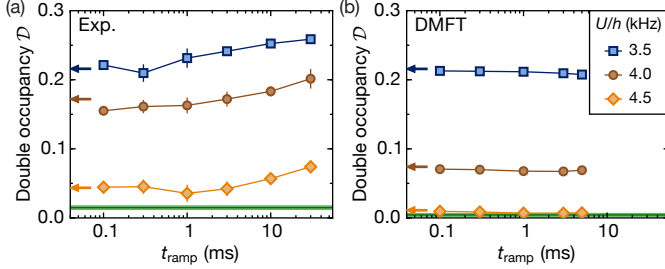

FIG. S13. Adiabaticity in near-resonant modulation. Experimentally (a) and theoretically (b) obtained data measuring double occupancies $\mathcal{D}$ in a driven system ($\Omega/2\pi = 3.5$ kHz) after ramping up K_0 from 0 to 1.44(2) within 10 ms, hold for 100 μs and then ramp K_0 down to 0 in a strongly-interacting configuration with $J_{x,y,z}/h = [190(34), 100(7), 100(7)]$ Hz and $U/h = 3521(65)$ (blue squares), 4026(77) (brown circles) and 4531(90) Hz (orange diamonds). For each of these protocols, a reference level of $\mathcal{D}$ (blue, brown and orange arrows) is taken right after the 100 μs hold time. A reference level of $\mathcal{D}$ in the strongly-interacting static system is also shown (green line). No decay of $\mathcal{D}$ is observed. Error bars are shown for experimental data indicating a standard error for 5 measurements.

with this ramp protocol, reaching the Floquet state on the timescales considered in this study is an irreversible process [S7]. The DMFT simulations reproduce this nonadiabaticity for the same protocol, which makes it clear that the irreversibility is inherent to the driven system and not related to experimental imperfections.

In a second set of measurements a two-step ramp protocol is performed (Fig. S14), in the same way as it was done in [S8]. Here we prepare the system at interaction $U_{\text{load}}/h = 5.0(1)$ kHz and then switch-on the drive by ramping up K_0 from 0 to 1.43(2) within 10 ms. Afterwards, while driving, we ramp down the interaction to
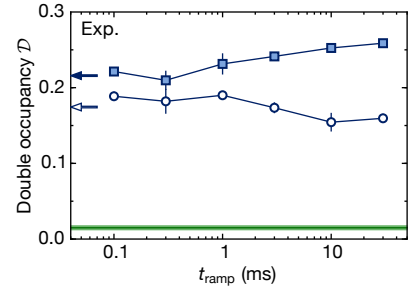

FIG. S14. Adiabaticity in near-resonant modulation for different ramp protocols. In a lattice with $J_{x,y,z}/h = [200(50), 99(9), 99(11)]$ Hz and $U/h = 3.5(1)$ kHz the modulation ($\Omega/(2\pi) = 3.5$ kHz, $K_0 = 1.44(3)$) is ramped in 10 ms. We hold another 100 μs before $\mathcal{D}$ is measured for different ramp down times. We compare a two-step ramping scheme (empty circles) with a single-step ramping scheme (squares) which was shown already in Fig. 4. In the two-step protocol the modulation is ramped up and down first at $U_{\text{load}}/h = 5.0(1)$ kHz away from resonance followed by a ramp of the interaction to $U_{\text{final}}/h = 3.5$ kHz and back respectively. Reference levels of $\mathcal{D}$ before ramping back are shown (arrows) as well as in the undriven lattice (solid line). Data points are the mean and standard error of 5 individual runs.

$U_{\text{final}}/h = 3.5(1)$ kHz in another 10 ms. Then, the interaction is ramped back to U_{load} within a variable amount of time t_{ramp}. Finally, K_0 is ramped down to 0 within another 10 ms before the detection of $\mathcal{D}$ starts. In addition, a reference level of $\mathcal{D}$ is taken right after the ramp down of the interaction from U_{final} to U_{load} to compare the amount of $\mathcal{D}$ that is adiabatically reverted with this protocol. In direct comparison to the single ramp protocol it can be seen that $\mathcal{D}$ decreases by around 10%. This relative decrease is very similar to the one seen in the theory data which was only taken for half filling (see Fig. S8).

H. Long-term dynamics in near-resonant modulation

We now comment on the experimental data for long-term dynamics $t_{\text{ramp}} \geq 10$ ms in the case of near-resonant driving, see Fig. 3. For the results with $U \approx \hbar\Omega$ and non-zero K_0 a decrease of double occupancy $\mathcal{D}$ is observed and no theoretical simulation comparison is provided in these cases. The underlying harmonic confinement can be described by a mean trapping frequency of $\overline{\omega}/(2\pi) = 86$ Hz (see Table V) which gives us an estimation on the timescale $t_{\text{trap}} \simeq 10$ ms above which trap induced dynamics can occur. Energy absorption from the drive can also lead to excitation of atoms to higher bands and shows up as unwanted heating in the lowest band models. An experimental indication of this heating is atom loss detected because atoms in higher bands can escape the trapped system. We only detect a significant atom loss on the order of 15% at the last point for 300 ms

Adiabaticity in near-resonant modulation

parameter	value(s)
ramp up K_0 protocol	
$J_{x,y,z}/h$ (Hz)	190(34), 100(7), 100(7)
W/h (Hz)	1178(79)
U/h (Hz)	3521(65), 4026(77), 4531(90)
K_0	1.44(2)
ramp up U and K_0 protocol	
$J_{x,y,z}/h$ (Hz)	211(38), 97(6), 97(8)
W/h (Hz)	1201(86)
U_{load}/h (Hz)	5032(106)
U_{final}/h (Hz)	3518(61)
K_0	1.43(2)

TABLE IV. Experimental parameters for the near-resonant modulation protocols to measure $\mathcal{D}$ versus either K_0 or U^{eff}. Each parameter value and error were computed as described in Table I.
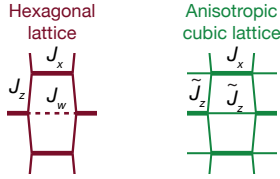

FIG. S15. Illustrations of different lattice configurations. In the actual experiment, a hexagonal lattice configuration is used (left). To estimate the entropy and density profile in the harmonic trap, a high-temperature series calculation is performed for an anisotropic cubic lattice (right). To match the bandwidths of the two configurations, we use $\tilde{J}_z = (2J_z + J_w)/3$.

ramp time. On the same very long timescale for this experiment we also expect technical heating arising mainly from fluctuations in the lattice depth.

I. Density profiles

In the experiment, the atoms are confined in a harmonic trapping potential. As a consequence, the density profile is inhomogeneous and all observables are averaged over the trap. In order to capture this in the numerical calculation, we determine the density profiles for our experimental parameters.

The main difficulty from an experimental point of view is to estimate the entropy of the atoms in the system. In order to determine it, we take additional reference measurements for two lattice configurations which are similar to the ones used in the off- and near-resonant driving regimes. For these systems, we measure three different observables:

- The double occupancy $\mathcal{D} = 2\langle\hat{n}_{i\uparrow}\hat{n}_{i\downarrow}\rangle$

- The transverse spin-spin correlator between neighboring sites in the x direction $\mathcal{C} = -\langle\hat{S}_i^x\hat{S}_{i+1}^x\rangle - \langle\hat{S}_i^y\hat{S}_{i+1}^y\rangle$

- The nearest-neighbor correlator $\mathcal{N} = \langle\hat{n}_i\hat{n}_{i+1}\rangle$

($\langle...\rangle$ denotes at the same time the expectation value and the trap average).

Afterwards, we run calculations based on a high-temperature series expansion [S9] and try to reproduce the measured quantities by adjusting the entropy S as a free parameter. All other parameters (Hubbard parameters, atom number and trapping frequencies) were derived from independent calibrations. The high-temperature series code is assuming an anisotropic cubic lattice geometry, which is very close to our hexagonal lattice configuration with the only difference that tunneling is allowed across the hexagonal unit cell (see Fig. S15). From this procedure, we obtain an estimate of the entropy per particle of $S/N = 2.2(4)$ k_B and $S/N = 2.4(2)$ k_B for the lattice configurations used in the off- and near-resonant driving regimes, respectively.

We use this estimate of the entropies to calculate self-consistently the temperature T and chemical potential μ with the same high-temperature series expansion. For this, we use the actual lattice configurations and trapping frequencies set in the experimental measurements. The results are summarized in Table V. Finally, we use these parameters to calculate the density profile in the trap using a local density approximation $n(\mathbf{r}) \equiv n(\mu - V_{\text{Trap}}(\mathbf{r}), T) = n(\mu - m/2\sum\omega_i^2 x_i^2, T)$, where

Quantity	Unit	Off-resonant	Near-resonant
Atom number N	10^3	35	35
Entropy S/N	k_B	2.2 ± 0.4	2.4 ± 0.2
Tunnelings $(J_x, J_y, \tilde{J}_z)/h$	Hz	487,102,83	202,100,67
Interaction U/h	kHz	1.1	3.5
Trap frequencies $(\omega_x, \omega_y, \omega_z)/(2\pi)$	Hz	66,65,140	64,80,123
Temperature T	J_x	$1.21^{+0.44}_{-0.36}$	$3.31^{+0.57}_{-0.51}$
Chem. potential μ	J_x	$1.42^{-0.74}_{+0.48}$	$3.87^{-0.98}_{+0.80}$

TABLE V. Parameters used to calculate the density profiles shown in Fig. S16. The entropy was estimated from reference measurements in similar lattice configurations. The tight-binding parameters and trap frequencies in the off-resonant regime are calculated for the intermediate lattice configuration, which sets the density profile. In the near-resonant regime, the parameters are the ones from the final lattice configuration (see Table III). In this case, the tight-binding parameters and mean trapping frequencies of the intermediate lattice configuration are matched to the final lattice. The upper and lower bounds on T and μ correspond to the maximum and minimum value of the entropy given above.
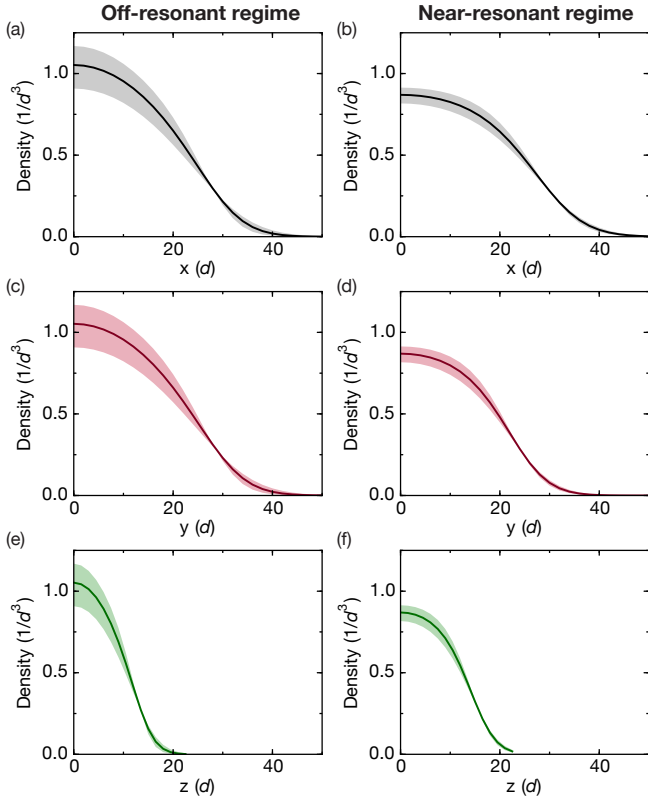
FIG. S16. Density profiles used for the numerical calculations. Cut through the density profiles starting from the center of the harmonic trap in the x-, y and z directions for the off-resonant (a, c, e) and near-resonant modulation regimes (b, d, f). The lattice spacing is denoted as d. The profiles were calculated with a high-temperature series expansion for the parameters given in Table V. The shading reflects the uncertainty of the density resulting from the error in temperature and chemical potential. This uncertainty is captured by the error bars shown for the numerical results in the figures of the main text. Both in the experiment and the numerical calculations the double occupancy is averaged over the inhomogeneous density profile of the cloud.

μ is the chemical potential in the center of the trap and $\omega_i/(2\pi)$ are the trapping frequencies in the x_i-direction. The corresponding profiles are shown in Fig. S16. All numerical results shown in the main text are averages of the double occupancy over these inhomogeneous densities, with the error bars representing the uncertainties in the temperature and chemical potential.

[S1] H. Aoki, N. Tsuji, M. Eckstein, M. Kollar, T. Oka, and P. Werner, Reviews of Modern Physics **86** (2014).

[S2] M. Eckstein and P. Werner, Physical Review B **82**, 115115 (2010).

[S3] A. Herrmann, Y. Murakami, M. Eckstein, and P. Werner, EPL (Europhysics Letters) **120**, 57001 (2017).

[S4] F. Peronaci, M. Schiró, and O. Parcollet, Physical Review Letters **120**, 197601 (2018).

[S5] M. Bukov, M. Kolodrubetz, and A. Polkovnikov, Physical Review Letters **116**, 125301 (2016).

[S6] T. Uehlinger, G. Jotzu, M. Messer, D. Greif, W. Hof-stetter, U. Bissbort, and T. Esslinger, Physical Review Letters **111**, 185307 (2013).

[S7] F. Görg, M. Messer, K. Sandholzer, G. Jotzu, R. Desbuquois, and T. Esslinger, Nature **553**, 481 (2018).

[S8] R. Desbuquois, M. Messer, F. Görg, K. Sandholzer, G. Jotzu, and T. Esslinger, Physical Review A **96**, 053602 (2017).

[S9] J. Oitmaa, C. Hamer, and W. Zheng, *Series Expansion Methods for Strongly Interacting Lattice Models* (Cambridge University Press, Cambridge, England, 2006).